CHEMICAL
RESEARCH,
-DEVELOPMENT &
ENGINEERING
CENTER

CRDEC-CR-120

IMPLEMENTATION SCHEME FOR RECURSION IN SPECTRAL DIMENSION

Chein-I Chang

UNIVERSITY OF MARYLAND
Baltimore, MD 21228-5398

March 1992

U.S. ARMY
ARMAMENT
MUNITIONS
CHEMICAL COMMAND

Aberdeen Proving Ground, Maryland 21010-5423

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1992 March	Final, 90 Sep - 91 Mar

4. TITLE AND SUBTITLE
Implementation Scheme for Recursion in Spectral Dimension

5. FUNDING NUMBERS
C-DAAD05-90-P-8392

6. AUTHOR(S)
Chang, Chein-I.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Electrical Engineering
University of Maryland
Baltimore County ~~Campus~~
Baltimore, MD 21228-5398

8. PERFORMING ORGANIZATION REPORT NUMBER
CRDEC-CR-120

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
CDR, CRDEC, ATTN: SMCCR-DDT, APG, MD 21010-5423

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
COR: Mark Althouse, SMCCR-DDT, (301) 671-5561

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Developing recursive detection algorithms for thermal image processing is very important for real time implementation. In his previous work for CRDEC, Warren used a first order autoregressive time series to model the background of thermal image with which he further developed a recursive algorithm in time frames. In this report, we revisit his work and present an alternative approach to developing a recursive algorithm for thermal image target detection. The new approach uses Kalman filter theory, which has proven to be very powerful in real time processing because of its recursive nature. The algorithm developed by Warren is designed for a single spectral band. In reality, however, the characteristics of the vapor cloud and background may vary from band to band in spectral domain; thus, it is practical to extend Warren's and the Kalman filtering approaches to cover multiple spectral bands. To alleviate the difficulty of processing multiple bands, two suboptimal models (separable spectral correlation and separable spectral-Markov spatial correlation) are also proposed for the background.

14. SUBJECT TERMS		15. NUMBER OF PAGES
Kalman filter Detection algorithm	Thermal imaging	37
		16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

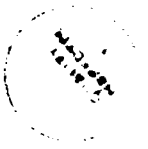
Blank

PREFACE

The work described in this report was authorized under Contract No. DAAD05-90-P-8392. This work was started in September 1990 and completed in March 1991.

The use of trade names or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.

This report has been approved for release to the public.



Blank

CONTENTS

Blank

IMPLEMENTATION SCHEME FOR RECURSION IN SPECTRAL DIMENSION

I. Introduction

One of major difficulties for thermal imaging processing is background suppression. Many different approaches have been proposed over the past years ranging from spatial domain analysis to frequency domain analysis. Recently consideration of coupling temporal correlation with either spatial or spectral correlation is also investigated to combat the degradation caused by poor signal to noise ratio or low concentrated clouds against background.

In [1,2], Warren and his colleagues proposed a mixed approach to process spatial images in frequency domain by using a first order autoregressive time series to account for temporal correlation. As a result, Warren developed a recursive algorithm for detectors to improve detection performance of low constrast signal to background clutter ratio. Although Warren also mentioned a possible approach by using a Kalman filter, he never discussed it. In this report we adopt a Kalman filter approach to derive similar results to that obtained by using an autoregressive model. This approach is more intuitive than Warren's approach because recursion is the nature of a Kalman filter.

It is known that Kalman fiter has a wide range of various applications. The most important feature Kalman filter possesses is recursion which can be implemented in real time processing. Because of its nature of updating information as time goes along it is worthwhile to consider Kalman filtering approach as opposed to Warren's approach. The main idea of Kalman filter is to introduce a new process, which is called innovation process, generated by the observation model. The great advantages of using the innovation process is orthogonality. The principle of orthogonality decorrelates observations obtained from different time frames. In the case of a Gaussian process, the induced innovation process becomes an independent process of time. As a consequence, if a thermal image detection problem can be transferred to an equivalent detection problem in terms of innovation processes, using the independency of the innovation process immediately will significantly reduce computational complexity and tremendously simplify mathematical derivations. This is demonstrated by comparing the Kalman filter approach to Warren's work. The use of the innovation processes enables

us to develop a recursive detection algorithm. Since the state and observation models used in Kalman filters are very general, the derived recursive algorithm will have more broader applications than does Warren's algorithm.

Both Warren's and Kalman filtering approaches are developed based on a single spectral band. However, in some practical cases, the target in one spectral band will have a large contrast with high altitude clouds and no contrast with low clouds, while in a second spectral band the target would have no contrast with high clouds but a large contrast with low-altitude clouds. Therefore, using two spectral bands can ensure a good contrast for any kind of cloud background. Because of advantages of using multiple bands the extension of these two approaches to multiple spectral bands is investigated. While it is useful to process multiple bands, there is an extra factor to be considered which is spectral correlation. In general, the dimension required for a detector to use multiple bands will be 4 times of that for a single band. To avoid computational load for real time processing two suboptimal models are proposed which are characterized by separable spectral correlation and separable spectral Markov spatial correlation respectively. These two model have been used in some applications in adpative filtering [3,4].

This report is organized as follows. Section II reviews the signal model used in thermal imaging processing and Section III is devoted to the target model. Both were developed by Warren [1,2]. Section IV rederives Warren's recursive detection algorithm which will serve as a basis for extension to multiple spectral bands. Section V presents a new approach to developing a similar recursive algorithm to Warren's algorithm by using well-knowing Kalman filter theory. Section VI is to extend results obtained in Sections IV and V to multiple spectral bands, particularly emphasizing two spectral bands. Section VII is a conclusion section which suggests some directions for future investigation and study.

II. Signal Model

The signal model to be considered in this report is suggested by Warren [1]. In this section we follow his approach and describe the derivation.

The optical power collected on the thermal detector plane at pixel location $\mathbf{x}_i = (x_i, y_i)$, $P_{d_j}(\mathbf{x}_i, t_k)$ can be written as the convolution of the entrance pupil signal power $P_{s_j}(\theta, t_k)$ and system point spread function PSF:

$$P_{d_j}(\mathbf{x}_i, t_k) = \int P_{s_j}(\theta, t_k) PSF(x_i/f - \theta) d\theta^2 + P_{n_j}(\mathbf{x}_i, t_k) \quad (1)$$

Here $P_d(\mathbf{x}_i, t_k)$ represents a time series of images labeled by index k and multiple spectral bands labeled by index j and is a generalization of a single spectral filter and single time expression:

$$P_d(\mathbf{x}_i) = \int P_s(\theta) PSF(\mathbf{x}_i/f - \theta) d\theta + P_n(\mathbf{x}_i)$$

where $\mathbf{x}_i$ denotes the pixel location in the detector plane, f is the receiver effective focal length.

The PSF models the combined effects of the thermal sensor on the spatial resolution of the images which includes terms that model the optical components of the system, finite detector size, and effects associated with scanning, electronics, and time response of the detector. Although in general the PSF depends upon spectral band, for well-designed systems operating in the 8-12 μm band such as TICM-II, the image quality is relatively uniform with wavelength. $P_{n_j}(\mathbf{x}_i, t_k)$ is additive, zero-mean detector noise in spectral band j, time k and is uncorellated with the signal. The entrance pupil signal power $P_{s_j}(\theta, t_k)$ labeled by line-of-sight direction vector θ is given by

$$P_{s_j}(\theta, t_k) = A_R \Omega_R \epsilon_0 \int_0^\infty \{F_j(\lambda)[B_\lambda(T_C) + \tau_A(\lambda) e^{-\rho_\lambda C L(\theta, t_k)} (B_\lambda(T_B(\theta, t_k)) - B_\lambda(T_C))]\} \quad (2)$$

in terms of the Planck function at temperature T and wavelength λ:

$$B_\lambda(T) \equiv \frac{2hc^2}{\lambda^5} [e^{(hc/\lambda kT)} - 1]^{-1}.$$

Because both the detector noise P_n and the clutter background temperature field $T_B(\theta)$ are random variables, so is the detector power P_d and must be characterized statistically. Fo the vapor cloud image enhancement/detection algorithm, only the first and second statistics of P_d are of interest. Denoting ensemble average by bar, the first-order statistic, mean can be given either by $E[P_{d_j}(\mathbf{x}_i, t_k)]$ or $\overline{P_{d_j}(\mathbf{x}_i, t_k)}$, and the second order statistc, covariance matrix $\Lambda_{P_{d_{jj'}}}(\mathbf{x}_i, t_k, \mathbf{x}_{j'}, t_{k'})$ given by

$$\Lambda_{P_{d_{jj'}}}(\mathbf{x}_i, t_k, \mathbf{x}_{i'}, t_{k'}) \equiv E\{[P_{d_j}(\mathbf{x}_i, t_k) - \overline{P_{d_j}(\mathbf{x}_i, t_k)}][P_{d_{j'}}(\mathbf{x}_{i'}, t_{k'}) - \overline{P_{d_{j'}}(\mathbf{x}_{i'}, t_{k'})}]\}.$$

In order to evaluate these moments we assume that the clutter temperature field $T_B(\theta, t_k)$ to be spatially wide-sense stationary, i.e., $\overline{T_B(\theta, t_k)} = T_B$ and

$$T_{B_{k,k'}}(\theta, \theta') \equiv E\{[T_B(\theta, t_k) - T_B][T_B(\theta', t_{k'}) - T_B]\}.$$

These conditions mean that the background fluctuations are uniform over the image field.

Evaluating the mean of P_d gives

$$\overline{P_d(\mathbf{x}_i)} = \int \overline{P_s(\theta, t_k)} PSF(\mathbf{x}_i/f - \theta) d\theta + P_n(\mathbf{x}_i) \tag{3}$$

with

$$\overline{P_{s_j}(\theta, t_k)} = A_R \Omega_R \epsilon_0 \int_0^\infty \{F_j(\lambda)[B_\lambda(T_C) + \tau_A(\lambda) e^{-\rho_\lambda C L(\theta, t_k)} (\overline{B_\lambda(T_B(\theta, t_k))} - B_\lambda(T_C))]\} \tag{4}$$

By expanding $B_\lambda(T_B(\theta, t_k))$ about $\overline{T_B}$

$$\overline{B_\lambda(T_B(\theta, t_k))} = B_\lambda(\overline{T_B}) + \frac{B''_\lambda(\overline{T_B})}{2} \sigma^2_{T_B} + O(\sigma^4_{T_B})$$

where $\sigma^2_{T_B}$ is the constant variance of the temperature background field, and B''_λ is the second derivative of the Planck function with respective T. Numerical evaluation of B''_{T_B} for $\overline{T_B}$ =290°K at $\lambda = 10\mu$m shows that $B''_\lambda / B_\lambda = 1.8 \times 10^{-4}$, so for $\sigma_{T_B} \leq 10$°K there is less than one percent error in approximating

$$\overline{B_\lambda(T_B(\theta, t_k))} \approx B_\lambda(\overline{T_B}). \tag{5}$$

Since

$$E[P_{s_j}(\theta, t_k) P_{n_{j'}}(\mathbf{x}_{j'}, t_{k'})] = 0,$$

$$\begin{aligned}\Lambda_{P_{d_{jj'}}}(\mathbf{x}_i, t_k; \mathbf{x}_{i'}, t_{k'}) &= \int \{PSF(\mathbf{x}_i/f - \theta)\Lambda_{s_{jj'}}(\theta, t_k; \theta', t_{k'}) \\ &\quad PSF(\mathbf{x}_{i'}/f - \theta)\} d\theta^2 d\theta'^2 + \Lambda_{n_{jj'}}(\theta, t_k; \theta', t_{k'}) \qquad (6)\end{aligned}$$

with

$$\begin{aligned}\Lambda_{s_{jj'}}(\theta, t_k; \theta', t_{k'}) = \quad & \{(A_R \Omega_R \epsilon_0)^2 \int F_j(\lambda) F_{j'}(\lambda') \tau_A(\lambda) \tau_A(\lambda') e^{[-\rho CL(\theta, t_k) - \rho CL(\theta', t_{k'})]}\} \times \\ & E\{[B_\lambda(T_B(\theta, t_k)) - \overline{T_B(\theta, t_k)}][B'_\lambda(T_B(\theta', t_{k'})) - \overline{T_B(\theta', t_{k'})}]\} d\lambda^2 d\lambda'^2\end{aligned}$$

By expanding $B_\lambda(T_B(\theta, t_k)$ around $\overline{T_B}$, i.e.,

$$B_\lambda(T_B(\theta, t_k))|_{\overline{T_B}} \approx \overline{T_B} + B'_\lambda(\overline{T_B})(T_B(\theta, t_k - \overline{T_B})$$

we obtain the following equalites:

$$\begin{aligned}& E\{[B_\lambda(T_B(\theta, t_k) - \overline{T_B(\theta, t_k)}][B'_\lambda(T_B(\theta', t_{k'}) - \overline{T_B(\theta', t_{k'})}]\} \\ &= E\{[B'_\lambda(\overline{T_B})(T_B(\theta, t_k) - \overline{T_B})][(B'_\lambda)(\overline{T_B})(T_B(\theta', t_{k'}) - \overline{T_B})]\} \\ &= B'_\lambda(\overline{T_B})[E\{[T_B(\theta, t_k) - \overline{T_B})][(T_B(\theta', t_{k'}) - \overline{T_B})]\} B'_{\lambda'}(\overline{T_B}) \qquad (7)\end{aligned}$$

where $B'_\lambda(\overline{T_B})$ is the first derivative of the Planck function evaluated at $\overline{T_B}$. If we denote

$$\Lambda_{T_{B_\lambda B_{\lambda'}}} \equiv E\{[T_B(\theta, t_k) - \overline{T_B})][(T_B(\theta', t_{k'}) - \overline{T_B})]\}.$$

Equation (7) can be rewritten as

$$B'_\lambda(\overline{T_B}) \Lambda_{T_{B_\lambda B_{\lambda'}}} B'_{\lambda'}(\overline{T_B}) \qquad (8)$$

which is a factored form of spectral and spatial components.

By taking advantage of equation (8)

$$\Lambda_{s_{jj'}}(\theta, t_k; \theta', t_{k'}) = G_j(\theta, t_k) \Lambda_{T_{B_\lambda B_{\lambda'}}} G_j(\theta', t_{k'}) \qquad (9)$$

with

$$G_j(\theta, t_k) \equiv A_R \Omega_R \epsilon_0 \int F_j(\lambda) \tau_A(\lambda) B'_\lambda(\overline{T_B}) e^{-\rho_\lambda C L(\theta, t_k)} d\lambda.$$

Finally, the signal model is derived as follows.

$$pdf(\mathbf{P}_d) = \frac{1}{\sqrt{2\pi^{N^2 ML} |\Lambda_{\mathbf{P}_d}|}} e^{[\frac{(\mathbf{P}_d - \mathcal{E}\{\mathbf{P}_d\})^T \Lambda_{\mathbf{P}_d}^{-1} (\mathbf{P}_d - \mathcal{E}\{\mathbf{P}_d\})}{2}]} \tag{10}$$

where the detector power has been expressed as a vector having components

$$(\mathbf{P}_d)_{ijk} \equiv P_{d_j}(\mathbf{x}_i, t_k)$$

and $i, 1 \leq i \leq N^2$, labels the detector plane pixel, $j, 1 \leq j \leq L$, labels the spectral band, and $k, 1 \leq k \leq M$, labels the frame time. The mean and covariance of $\mathbf{P}_d$ are given by equations (3) and (9) respectively.

III. Target Model

To construct a target model appropriate for image enhancement of vapor clouds against clutter backgrounds we rewrite equation (2) as

$$\overline{P_{s_j}}(\theta, t_k) = P_{A_j} + \overline{[P_{B_j}]}\overline{[e^{-\rho_\lambda CL(\theta,t_k)}]}_j$$

where

$$P_{A_j} = A_R \Omega_R \epsilon_0 \int F_j(\lambda) B_\lambda(T_C) d\lambda,$$

and

$$\overline{P_{B_j}} = A_R \Omega_R \epsilon_0 \int F_j(\lambda) \tau_A(\lambda) \Delta B_\lambda d\lambda,$$

with $\Delta B_\lambda = B_\lambda(\overline{T_B}) - B_\lambda(T_C)$, and a vapor component defined as

$$\overline{[e^{-\rho_\lambda CL(\theta,t_k)}]}_j \equiv \frac{\int F_j(\lambda)\tau_A(\lambda)[e^{-\rho_\lambda CL(\theta,t_k)}]\Delta B_\lambda d\lambda}{\int F_j(\lambda)\tau_A(\lambda)\Delta B_\lambda d\lambda}.$$

The above equation can be interpreted as a spectral average over the jth spectral band of the transmission loss produced by the path-integrated concentration $CL(\theta, t_k)$ along the line-of-sight θ at time t_k. For narrow band optical filters, τ_A and ΔB_λ are approximately constant over the support of the bandpass function F_j, giving

$$\overline{[e^{-\rho_\lambda CL(\theta,t_k)}]}_j \equiv \frac{1}{\Delta\lambda_j} \int F_j(\lambda)[e^{-\rho_\lambda CL(\theta,t_k)}] d\lambda$$

with $\Delta\lambda_j \equiv \int F_j(\lambda) d\lambda$ the effective bandpass width of the jth optical filter.

IV. Recursive Detection Algorithm Developed by Warren

The main result in Warren's work is to present a recursive detection algorithm to process thermal images in frequency domain. The idea is to model the background as a first order autoregressive time series by which the detection of vapor clouds can be done recursively in time frames. The approach is to use the Markov property provided by the autoregressive model to decompose a test statistic into two parts, a test statistic obtained from the preceeding time frame and observation generated at the present time. Because of Markov property, the past information only depends on the time frame immediately before the current time frame. Therefore, the latest test statistic already provides all the past information. By taking of this advantage Warren derived a recursive algorithm in time for real time processing. In this section, Warren's algorithm will be rederived. However, there is an alternative which has a more general aspect than does Warren's approach, that is Kalman filter approach which will be investigated in detail in the next section.

Image Model

The image model used in Warren's work is described by

$$\begin{aligned} H_0 : \quad & I_k(\mathbf{x}) = B_k(\mathbf{x}) + N_k(\mathbf{x}) \\ versus & \\ H_1 : \quad & I_k(\mathbf{x}) = T_k(\mathbf{x}) + B_k(\mathbf{x}) + N_k(\mathbf{x}) \end{aligned} \tag{11}$$

where the background is modeled by a first order autoregressive time series

$$B_k(\mathbf{x}) - \overline{B} = \sum_{\mathbf{x}'} G(x - x')[B_{k-1}(\mathbf{x}) - \overline{B}] + e_k(\mathbf{x}), \tag{12}$$

and their corresponding correlation matrices are

$$\Lambda_B(\mathbf{x}, \mathbf{x}')_{kk'} = E[[(B_k(\mathbf{x}) - \overline{B})(B_{k'}(\mathbf{x}) - \overline{B})] \tag{13}$$

$$
\begin{aligned}
\Lambda_e(\mathbf{x},\mathbf{x}')_{kk'} &= E[e_k(\mathbf{x})e_{k'}(\mathbf{x})] \\
&= [\Lambda_{Bkk'} - G \bigotimes \Lambda_{Bkk'} \bigotimes G](\mathbf{x},\mathbf{x}')\delta_{kk'} \qquad (14) \\
C_n(\mathbf{x}-\mathbf{x}')\delta_{kk'} &= E[(N_k(\mathbf{x}) - \overline{N})(N_k(\mathbf{x}) - \overline{N})] \\
T_k(\mathbf{x}) &= AS(\mathbf{x} - \mathbf{x}_T(k)) \\
E[I_k(\mathbf{x})|H_0] &= \overline{I}_0 = \overline{B} + \overline{N} \\
E[I_k(\mathbf{x})|H_1] &= \overline{I}_T = T_k(\mathbf{x}) + \overline{I}_0
\end{aligned}
$$

As mentioned in Warren's work, because of correlations between different time frames- the image model described above is not directly useful for developing detectors whose test statistics take the form of ratio of one likelihood function to another likelihood function. In order to circumvent the dilemma, Warren Fourier transferred samples in spatial domain to frequency domain with which statitsical and computational advantages can be gained.

Algorithms Developed in Frequency Domain

$$
\begin{aligned}
a_k(\mathbf{u}) &= \sum_{\mathbf{x}} [I_k(\mathbf{x}) - \overline{I}_0] e^{-i2\pi \mathbf{u}\cdot\mathbf{x}} \\
b_k(\mathbf{u}) &= \sum_{\mathbf{x}} [B_k(\mathbf{x}) - \overline{B}] e^{-i2\pi \mathbf{u}\cdot\mathbf{x}} \\
n_k(\mathbf{u}) &= \sum_{\mathbf{x}} [N_k(\mathbf{x}) - \overline{N}] e^{-i2\pi \mathbf{u}\cdot\mathbf{x}} \\
t_k(\mathbf{u}) &= \sum_{\mathbf{x}} [T_k(\mathbf{x})] e^{-i2\pi \mathbf{u}\cdot\mathbf{x}} = A e^{-i2\pi \mathbf{u}\cdot\mathbf{x}} \mathbf{S}(\mathbf{u})
\end{aligned}
$$

Based on the above discrete Fourier transforms, an equivalent test hypothesis problem to that described by (11) is given as follows.

$$
\begin{aligned}
H_0 : &\quad a_k(\mathbf{u}) = b_k(\mathbf{u}) + n_k(\mathbf{u}) \\
versus & \qquad\qquad (15) \\
H_1 : &\quad a_k(\mathbf{u}) = t_k(\mathbf{u}) + b_k(\mathbf{u}) + n_k(\mathbf{u})
\end{aligned}
$$

where we assume $E[a_k(\mathbf{u})|H_0] = 0$ and $E[a_k(\mathbf{u})|H_1] = t_k(\mathbf{u})$. Under mild conditions, the correlation matrix of $a_k(\mathbf{u})$ betwen $a_{k'}(\mathbf{u})$ is given by

$$\Lambda_a(\mathbf{u}, \mathbf{u}')_{kk'} = E[b_k^H(\mathbf{u})b_{k'}(\mathbf{u}')] + E[n_k^H(\mathbf{u})n_{k'}(\mathbf{u}')]$$

where H is a Hermitian conjugation. which results in

$$\Lambda_a(\mathbf{u}, \mathbf{u}')_{kk'} = \{E[b_k^H(\mathbf{u})b_{k'}(\mathbf{u}')] + E[|n_k(\mathbf{u})|^2]\delta_{kk'}\}\delta^2(\mathbf{u} - \mathbf{u}')] \tag{16}$$

Since the backgorund is assumed to be a first order autoregressive model, the corresponding discrete Fourier transforms (DFT) in frequency domain is given by

$$b_k(\mathbf{u}) = \gamma(\mathbf{u})b_{k-1}(\mathbf{u}) + \beta_k(\mathbf{u}) \tag{17}$$

where

$$\gamma(\mathbf{u}) = \sum_{\mathbf{x}} G(\mathbf{x})e^{-i2\pi\mathbf{u}\cdot\mathbf{x}} \tag{18}$$

$$\beta(\mathbf{u}) = \sum_{\mathbf{x}} e(\mathbf{x})e^{-i2\pi\mathbf{u}\cdot\mathbf{x}} \tag{19}$$

Let $\Phi_b(\mathbf{u})$ and $\Phi_n(\mathbf{u})$ be corresponding DFTs of $\Lambda_b(\mathbf{x})$ and $C_n(\mathbf{x})$. Then

$$\Phi_b(\mathbf{u}) = \sum_{\mathbf{x}} \Lambda_b(\mathbf{x})e^{-i2\pi\mathbf{u}\cdot\mathbf{x}} \tag{20}$$

$$\Phi_n(\mathbf{u}) = \sum_{\mathbf{x}} C_n(\mathbf{x})e^{-i2\pi\mathbf{u}\cdot\mathbf{x}} \tag{21}$$

with

$$E[\beta_k(\mathbf{u})] = 0, \tag{22}$$

$$E[\beta_k(\mathbf{u})\beta_{k'}(\mathbf{u}')] = \sigma_\beta^2(\mathbf{u})\delta^2(\mathbf{u} - \mathbf{u}')\delta_{kk'} \tag{23}$$

$$\sigma_\beta^2(\mathbf{u}) = [1 - |\gamma(\mathbf{u})|^2]\Phi_b(\mathbf{u}). \tag{24}$$

Combining the above equations yields, for $k > k'$

$$\Lambda_a(\mathbf{u})_{kk'} = \{[\gamma(\mathbf{u})]^{k-k'}\Phi_b(\mathbf{u}) + \Phi_n(\mathbf{u})\delta_{kk'}\}\delta_{kk'}(\mathbf{u}). \tag{25}$$

As mentioned in Warren's work, under mild conditions the probability density function of the corresponding DFT is also a Gaussian distribution under the hypothesis H_i and given by

$$\mathbf{p}(\mathbf{a}(M)|H_i) = \frac{1}{\pi^{Mn^2}|\mathbf{\Lambda_a}(M)|}\exp\{-\frac{1}{2}[\sum_{k,k'=1}^{M,M}\sum_{\mathbf{u}}(a_k(\mathbf{u})-[\overline{a_{k|i}(\mathbf{u})}])^H\Lambda_a^{-1}(\mathbf{u})_{kk'}[a_{k'}(\mathbf{u})-\overline{a_{k|i}(\mathbf{u})}]]\}$$

where we define

$$\begin{aligned}\overline{a_{k|i}(\mathbf{u})} &\equiv E[a_k(\mathbf{u})|H_i] \\ \overline{a_{k'|i}(\mathbf{u})} &\equiv E[a_{k'}(\mathbf{u})|H_i]\end{aligned}$$

and $\mathbf{\Lambda_a}(M)$ is the autocovariance matrix of the random vector $[a_k(\mathbf{u}), a_{k-1}(\mathbf{u}), \cdots, a_1(\mathbf{u})]$.

By defining

$$\begin{aligned}\mathbf{a}(k) &\equiv \mathbf{a}_k(\mathbf{u}) \equiv [a_k(\mathbf{u}), a_{k-1}(\mathbf{u}), \cdots, a_1(\mathbf{u})]^T \\ \mathbf{a}(k+1) &\equiv \mathbf{a}_{k+1}(\mathbf{u}) \equiv [a_{k+1}(\mathbf{u}), \mathbf{a}_k(\mathbf{u})]^T\end{aligned}$$

we obtain

$$||\mathbf{a}(M) - \overline{\mathbf{a}(M|i)}||^2_{\mathbf{\Lambda_a}^{-1}(M)} \equiv \sum_{k,k'=1}^{M,M}\sum_{\mathbf{u}}(a_k(\mathbf{u}) - \overline{a_{k|i}}(\mathbf{u}))^H\Lambda_a^{-1}(\mathbf{u})_{kk'}(a_{k'}(\mathbf{u}) - \overline{a_{k'|i}(\mathbf{u})}). \quad (26)$$

Taking logorithm of ratio of $\mathbf{p}(\mathbf{a}(M)|H_1)$ to $\mathbf{p}(\mathbf{a}(M)|H_0)$ yields the log likelihood ratio test given by

$$LR_{\mathbf{a}(M)}(A, \mathbf{x}_T) = log\ \frac{\mathbf{p}(\mathbf{a}(M)|H_1)}{\mathbf{p}(\mathbf{a}(M)|H_0)}. \quad (27)$$

The log likelihood ratio test $LR_{\mathbf{a}(M)}(A, \mathbf{x}_T)$ obtained from the above equation is the desired detector to be used for thermal image detection. Unfortunately, the DFT of the autocovariance matrix generated by $\mathbf{a}(k)$, $\mathbf{\Lambda_a}(k)$ becomes very large as time frames k are accumulated. In particular, calculating the inverse of $\mathbf{\Lambda_a}(k)$ becomes extremely complicated. One way to avoid this difficulty is to compute $LR_{\mathbf{a}(k)}(A, \mathbf{x}_T)$ recursively until k reaches M.

In what follows, we follow Warren's approach to derive a recursive formula for updating the log likelihood ratio test when time moves along.

Recursive Formula

First of all, we express the autocovariance matrix of $\mathbf{a}(k+1)(\mathbf{u})$ at time $k+1$ in terms of that of $\mathbf{a}(k)(\mathbf{u})$ at time k,

$$\mathbf{\Lambda_a}(k+1) = \begin{bmatrix} \Delta_0 & \underline{D}_k^H \\ \underline{D}_k & \mathbf{\Lambda_a}(k) \end{bmatrix} \tag{28}$$

where

$$\begin{aligned} \Delta_0 &\equiv \Phi_b + \Phi_n \\ \underline{D}_k^H &= \Phi_b[\gamma^H, (\gamma^2)^H, \cdots, (\gamma^k)^H] \\ \underline{D}_{k+1}^H &= \gamma^H[\Phi_b, \underline{D}_k^H] \end{aligned}$$

In order to invert the matrix we use the following matrix identities in matrix algebra. Assume that the matrix Ξ has the following form:

$$\Xi = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Then the inverse of Ξ can be calculated as follows.

$$\begin{aligned} \Xi^{-1} &= \begin{bmatrix} A^{-1} + A^{-1}D(B - CA^{-1}D)^{-1}CA^{-1} & -A^{-1}D(B - CA^{-1}D)^{-1} \\ -(B - CA^{-1}D)^{-1}CA^{-1} & (B - CA^{-1}D)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (A - DB^{-1}C)^{-1} & -(A - DB^{-1}C)^{-1}DB^{-1} \\ -B^{-1}C(A - DB^{-1}C)^{-1} & B^{-1} + B^{-1}C(A - DB^{-1}C)^{-1}DB^{-1} \end{bmatrix}. \end{aligned} \tag{29}$$

Applying the above matrix identities (29) to $\mathbf{\Lambda_a}(k+1)$ yields

$$\mathbf{\Lambda_a}^{-1}(k+1) = \begin{bmatrix} \Delta_k^{-1} & -\Delta_k^{-1}\mathbf{Y}_k^H \\ -\mathbf{Y}_k\Delta_k^{-1} & \mathbf{\Lambda_a}^{-1}(k) + \mathbf{Y}_k\Delta_k^{-1}\mathbf{Y}_k^H \end{bmatrix}. \tag{30}$$

where

$$\Delta_k = \Delta_0 - \underline{D}_k^H\mathbf{Y}_k \tag{31}$$

$$\mathbf{Y}_k \equiv \mathbf{\Lambda_a}^{-1}(k)\underline{D}_k \tag{32}$$

$$\Delta_{k+1} = \Phi_b(1-|\gamma|^2) + \Phi_n(1+|\gamma|^2) - |\gamma|^2\Phi_n\Delta_k^{-1}\Phi_n \tag{33}$$

$$\mathbf{Y}_{k+1}^H = \gamma^H[1-\Phi_n\Delta_k^{-1}, \Phi_n\Delta_k^{-1}\mathbf{Y}_k^H] \tag{34}$$

Now if we define

$$\overline{\mathbf{a}(k|i)} \equiv E[\mathbf{a}(k)|H_i]$$

$$\overline{\mathbf{a}(k+1|i)} \equiv E[\mathbf{a}(k+1)|H_i]$$

and substituting equation (30) into equation (26) yields

$$||\mathbf{a}(k+1) - \overline{\mathbf{a}(k+1|i)}||^2_{\Lambda_a^{-1}(k+1)} = ||\mathbf{a}(k) - \overline{\mathbf{a}(k|i)}||^2_{\Lambda_a^{-1}(k)} + ||a_{k+1} - a_{k+1|i}||^2_{\Delta_k^{-1}},$$

and

$$|\mathbf{\Lambda_a}(k+1)| = |\Delta_k|\ |\mathbf{\Lambda_a}(k)|.$$

Using

$$\mathbf{p}(\mathbf{a}(k+1)|H_i) = \mathbf{p}(a_{k+1}|H_i, \mathbf{a}(k))\mathbf{p}(\mathbf{a}(k)|H_i) \tag{35}$$

we derive the following conditional probability density function given that $H_i, \mathbf{a}(k)$.

$$\mathbf{p}(a_{k+1}|H_i, \mathbf{a}(k)) = \frac{\mathbf{p}(\mathbf{a}(k+1)|H_i)}{\mathbf{p}(\mathbf{a}(k)|H_i)} \tag{36}$$

$$= \frac{1}{\pi^{N^2}|\Delta_k|}\exp\{-\frac{1}{2}||a_{k+1} - \hat{a}_{k+1|k,i}||^2_{\Delta_k^{-1}}\} \tag{37}$$

where

$$\hat{a}_{k+1|k,i} \equiv \overline{a_{k+1|i}} + \mathbf{Y}_k^H\{\mathbf{a}(k) - \overline{\mathbf{a}(k|i)}\} \tag{38}$$

$$= \overline{a_{k+1|i}} + \gamma^H\{(1-\Phi_n\Delta_{k-1}^{-1})(a_k(\mathbf{u}) - \overline{\mathbf{a}(k|i)}\} + \Phi_n\Delta_{k-1}^{-1}[\hat{a}_{k|k-1,i}] - \overline{\mathbf{a}(k+1|i)} \tag{39}$$

Plugging equation (37) into (35) results in a desired recursive formula for log likelihood ratio test $LR_{\mathbf{a}(k+1)}(A, \mathbf{x}_T)$ in terms of $LR_{\mathbf{a}(k)}(A, \mathbf{x}_T)$.

$$LR_{\mathbf{a}(k+1)}(A,\mathbf{x}_T) = LR_{\mathbf{a}(k)}(A,\mathbf{x}_T) - \frac{1}{2}||a_{k+1|1} - \hat{a}_{k+1|k,1}||^2_{\Delta_k^{-1}} + \frac{1}{2}||a_{k+1|0} - \hat{a}_{k+1|k,0}||^2_{\Delta_k^{-1}} \quad (40)$$

Using equations (38-39) to interpret (15), we obtain

$$\hat{a}_{k+1|k,0} = \mathbf{Y}_k^H \mathbf{a}(k) \quad (41)$$

$$\hat{a}_{k+1|k,1} = t_{k+1} + \mathbf{Y}_k^H(\mathbf{a}(k) - \mathbf{t}(k)). \quad (42)$$

Substituting equations (41-42) we further obtain

$$LR_{\mathbf{a}(k+1)}(A,\mathbf{x}_T) = LR_{\mathbf{a}(k)}(A,\mathbf{x}_T) - \frac{1}{2}||t_{k+1} - \mathbf{Y}_k^H\mathbf{t}(k)||^2_{\Delta_k^{-1}} + [t_{k+1} - \mathbf{Y}_k^T\mathbf{t}(k)]^H[\Delta_k^{-1}][a_{k+1} - \mathbf{Y}^H\mathbf{a}(k)] \quad (43)$$

Finally, if we define

$$\hat{t}_{k+1|k} \equiv \mathbf{Y}^H\mathbf{t}(k) = \gamma^H[(1 - \Phi_n\Delta_{k-1}^{-1})t_k + \Phi_n\Delta_{k-1}^{-1}\hat{t}_{k|k-1}]$$

then a nice recursive form can be derived and given by

$$LR_{\mathbf{a}(k+1)}(A,\mathbf{x}_T) = LR_{\mathbf{a}(k)}(A,\mathbf{x}_T) - \frac{1}{2}||t_{k+1} - \hat{t}_{k+1|k}||^2_{\Delta_k^{-1}} \quad (44)$$

$$+[t_{k+1} - \hat{t}_{k+1|k}]^H[\Delta_k^{-1}][a_{k+1} - \hat{a}_{k+1|k,0}] \quad (45)$$

where

$$\hat{a}_{k+1|k,0} = \gamma^H[(1 - \Phi_n\Delta_{k-1}^{-1})a_k + \Phi_n\Delta_{k-1}^{-1}(\hat{a}_{k|k-1,0})] \quad (46)$$

with initial conditions

$$LR_{\mathbf{a}(0)}(A,\mathbf{x}_T) = 0,\ \hat{t}_{1|0} = 0,\ \hat{a}_{1|0,i} = 0$$

V. Recursive Detection Algorithm Using Kalman Filtering Approach

The idea of using Kalman filter theory is obvious because of recursion. Unlike Warren's work, this approach can be used to model background of different types. Recall that the background in Warren's work was assumed to a first order autoregressive (AR) model. As a result, Warren could use the Markov property induced by the AR model to derive a recursive formula for detectors described by a sequential likelihood ratio test statistic. In what follows, we will still adopt Warren's model, instead, consider a different approach to derive a similar recursive formula for detectors. The approach to be presented is to use Kalman filter theory. The major advantage of using a Kalman filter over Warren's approach is that the formula to be derived can be extended to cover more general models to describe background.

The core of Kalman filter theory is to introduce a new process, innovation process suggested by Kailath. Instead of directly dealing with the observation process, the innovation process is generated by collecting and updating new information as time goes along. In other words, given an observation process it is not necessary to store all information available up to the processing time because some information will be useless and some will be repeatedly stored which wastes storage. A more efficient way to manage information is to store all necessary information only once and dump unnecessary or unwanted information. The innovation process is developed based on this need. In general an observation process can be decomposed into two processes, predicted process and unpredicted process. The predicted process contains all previous information required for processing and the unpredicted process presents new information available at the processing time but not contained in the predicted process. Such unpredicted process resulting from the observation process is generally referred to as an innovation process.

In Kalman filter theory two models need to be specified which are a **State Model** described by a state or process equation and an **Observation Model** described by a measurement equation. In order to apply Kalman filter theory we interpret the image model

used in Warren's work as follows.

Let the background $b_k(\mathbf{u})$ be a state model chracterized by the state equation or process equation under hypothesis H_i given by

$$b_k(\mathbf{u}) = \gamma(\mathbf{u})b_{k-1}(\mathbf{u}) + \beta_k(\mathbf{u}) \tag{47}$$

and the observation model $a_{k|i}(\mathbf{u})$ depicted by measurement equation given by

$$a_{k|i}(\mathbf{u}) = s_{k|i}(\mathbf{u}) + b_k(\mathbf{u}) + n_k(\mathbf{u}) \tag{48}$$

where $s_{k|0}(\mathbf{u}) = 0$ for H_0, no target present and $s_{k|1}(\mathbf{u}) = t_1(\mathbf{u})$ for H_1, target present .

Since the derivation for a recursive Kalman filter under hypothesis H_1 is exactly the same as hypothesis H_0, we only consider the case for H_0 to simplify notations.

Because there are two models involved we define two innovation process, the unpredicted observation process $\alpha_k(\mathbf{u})$ derived from the measurement equation and the predicted state error process $\epsilon_{k+1,k}(\mathbf{u})$ derived from the state equation with their associated correlation matrices given by

$$\begin{aligned} \alpha_k(\mathbf{u}) &\equiv a_k(\mathbf{u}) - \hat{a}_{k|k-1}(\mathbf{u}) \\ \epsilon_{k,k-1}(\mathbf{u}) &\equiv b_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u}) \end{aligned}$$

$$\begin{aligned} \Phi_{\alpha,k}(\mathbf{u}) \equiv E[\alpha_k(\mathbf{u})\alpha_k^H(\mathbf{u})] &\equiv \sum_{\mathbf{x}} \Lambda_{\alpha,k}(\mathbf{x})e^{-2\pi\mathbf{ux}} \\ \Phi_{\epsilon,k,k-1}(\mathbf{u}) \equiv E[\epsilon_{k,k-1}(\mathbf{u})\epsilon_{k,k-1}^H(\mathbf{u})] &\equiv \sum_{\mathbf{x}} \Lambda_{\epsilon,k,k-1}(\mathbf{x})e^{-2\pi\mathbf{ux}} \end{aligned}$$

where the notation H is a Hermitian tranfomation, $\Lambda_{\alpha,k}(\mathbf{x})$ and $\Lambda_{\epsilon,k,k-1}(\mathbf{x})$ are correlation matrices of $\alpha_k(\mathbf{x})$ and $\epsilon_{k,k-1}(\mathbf{x})$ obtained in the spatial domain time $k, k-1$ respectively.

As we can see from above, the innovation process $\alpha_k(\mathbf{u})$ is the new information available which is contained in the process $a_k(\mathbf{u})$ observed at time k but not in $\hat{a}_{k|k-1}(\mathbf{u})$ where is the predicted process obtained from the past information up to time $k-1$. Similarly, $\epsilon_{k,k-1}(\mathbf{u})$

is the innovation process represents the new information provided by the state model $b_k(\mathbf{u})$ observed at time k and is obtained by substracting the predicted process $\hat{b}_{k|k-1}(\mathbf{u})$ from $b_k(\mathbf{u})$ so that the predicted information from $b_k(\mathbf{u})$ up to time $k-1$ can be removed. One of most important features that innovation processes possess is that they are white, i.e., independent processes. The property of independency makes problems extremely easy to deal with. The observation of such independency is intuitive. Since an innovation process contains unpredicted new information obtained at different time frames, the information obtained at a certain time frame must be independent of other time frames due to the nature of unpredictability. The relationship between these two innovation processes can be demonstrated by the following equation.

$$\begin{aligned} \alpha_k(\mathbf{u}) &= b_k(\mathbf{u}) + n_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u}) \\ &= \epsilon_{k,k-1}(\mathbf{u}) + n_k(\mathbf{u}) \end{aligned}$$

and

$$\begin{aligned} \Phi_{\alpha,k}(\mathbf{u}) &= E[\epsilon_{k,k-1}(\mathbf{u})\epsilon^H_{k,k-1}(\mathbf{u})] + \Phi_n(\mathbf{u}) \\ &= \Phi_{\epsilon,k,k-1}(\mathbf{u}) + \Phi_n(\mathbf{u}) \end{aligned}$$

Suppose that we fix time k and consider an estimate of $b_l(\mathbf{u})$ at time l, $\hat{b}_{l|k}(\mathbf{u})$ based on the information available up to time k. As we mentioned previously, the innovation process $\alpha_j(\mathbf{u})$ provides new information at time frame j. It is obvious that $\hat{b}_{l|k}(\mathbf{u})$ can be expressed in terms of a linear combination of $\alpha_j(\mathbf{u})$s from $j = 1$ to k .

$$\hat{b}_{l|k}(\mathbf{u}) = \sum_{j=1}^{k} D_{l,j}\alpha_j(\mathbf{u}). \tag{49}$$

In order to determine the coefficients $D_{l,j}$, we use the orthogonality principle which gives

$$E[\epsilon_{l,k}(\mathbf{u})\alpha^H_m(\mathbf{u})] = E[(b_l(\mathbf{u}) - \hat{b}_{l|k}(\mathbf{u}))\alpha^H_m(\mathbf{u})] = 0; \quad m = 1, 2, \cdots, k,$$

and results in

$$E[(b_l(\mathbf{u})\alpha^H_m(\mathbf{u})] = E[\hat{b}_{l|k}(\mathbf{u})\alpha^H_m(\mathbf{u})]$$

$$
\begin{aligned}
&= E[\{\sum_{j=1}^{k} D_{l,j}\alpha_j(\mathbf{u})\}\alpha_m^H(\mathbf{u})] \\
&= D_{l,m}\Phi_{\alpha,m}(\mathbf{u}).
\end{aligned}
$$

$D_{l,m}$ can be therefore solved by

$$
D_{l,m} = E[b_l(\mathbf{u})\alpha_m^H(\mathbf{u})]\Phi_{\alpha,m}^{-1}(\mathbf{u}).
$$

Plugging the solved $D_{l,j}$ into equation (49) yields

$$
\hat{b}_{l|k}(\mathbf{u}) = \sum_{j=1}^{k}\{E[b_l(\mathbf{u})\alpha_j^H(\mathbf{u})]\Phi_{\alpha,j}^{-1}(\mathbf{u})\}\alpha_j(\mathbf{u}).
$$

Now if we let $l = k+1$,

$$
\begin{aligned}
\hat{b}_{k+1|k}(\mathbf{u}) &= \sum_{j=1}^{k}\{E[b_{k+1}(\mathbf{u})\alpha_j^H(\mathbf{u})]\Phi_{\alpha,j}^{-1}(\mathbf{u})\}\alpha_j(\mathbf{u}) \\
&= \sum_{j=1}^{k-1}\{E[b_{k+1}(\mathbf{u})\alpha_j^H(\mathbf{u})]\}\Phi_{\alpha,j}^{-1}(\mathbf{u})\alpha_j(\mathbf{u}) + E[b_{k+1}(\mathbf{u})\alpha_k^H(\mathbf{u})]\Phi_{\alpha,k}^{-1}(\mathbf{u})\alpha_k(\mathbf{u}) \\
&= \sum_{j=1}^{k-1} E[\{\gamma(\mathbf{u})b_k(\mathbf{u}) + \beta_k(\mathbf{u})\}\alpha_j^H(\mathbf{u})]\Phi_{\alpha,j}^{-1}(\mathbf{u})\alpha_j(\mathbf{u}) \\
&\quad + E[\{\gamma(\mathbf{u})b_k(\mathbf{u}) + \beta_k(\mathbf{u})\}\alpha_k^H(\mathbf{u})]\Phi_{\alpha,k}^{-1}(\mathbf{u})\alpha_k(\mathbf{u}) \\
&= \gamma(\mathbf{u})\{\sum_{j=1}^{k-1} E[b_k(\mathbf{u})\alpha_j^H(\mathbf{u})]\Phi_{\alpha,j}^{-1}(\mathbf{u})\alpha_j\} \\
&\quad + \gamma(\mathbf{u})\{E[b_k(\mathbf{u})\alpha_k^H(\mathbf{u})]\Phi_{\alpha,k}^{-1}(\mathbf{u})\alpha_k(\mathbf{u})\} \\
&= \gamma(\mathbf{u})\hat{b}_{k|k-1}(\mathbf{u}) + \gamma(\mathbf{u})\{E[b_k(\mathbf{u})\alpha_k^H(\mathbf{u})]\Phi_{\alpha,k}^{-1}(\mathbf{u})\alpha_k(\mathbf{u})\} \\
&= \gamma(\mathbf{u})\hat{b}_{k|k-1}(\mathbf{u}) + G_k(\mathbf{u})\alpha_k(\mathbf{u})
\end{aligned}
$$

where $G_k(\mathbf{u})$ is called the Kalman gain given by

$$
\begin{aligned}
G_k(\mathbf{u}) &= \gamma(\mathbf{u})\{E[b_k(\mathbf{u})\alpha_k^H(\mathbf{u})]\Phi_{\alpha,k}^{-1}(\mathbf{u})\} \\
&= \gamma(\mathbf{u})\{E[b_k(\mathbf{u})(b_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u}) + n_k(\mathbf{u}))^H]\Phi_{\alpha,k}^{-1}(\mathbf{u})\} \\
&= \gamma(\mathbf{u})\{E[b_k(\mathbf{u})(\epsilon_{k,k-1}(\mathbf{u}) + n_k(\mathbf{u}))^H]\Phi_{\alpha,k}^{-1}(\mathbf{u})\}.
\end{aligned}
$$

Since $E[\hat{b}_{k|k-1}(\mathbf{u})\epsilon^H_{k,k-1}(\mathbf{u})] = 0$, we can add it to the above equation and obtain

$$\begin{aligned} G_k(\mathbf{u}) &= \gamma(\mathbf{u})\{E[(b_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u}))(\epsilon^H_{k,k-1}(\mathbf{u}))]\Phi^{-1}_{\alpha,k}(\mathbf{u})\} && (50)\\ &= \gamma(\mathbf{u})\{E[\epsilon_{k,k-1}(\mathbf{u})\epsilon^H_{k,k-1}(\mathbf{u})]\Phi^{-1}_{\alpha,k}(\mathbf{u})\} && (51)\\ &= \gamma(\mathbf{u})\Phi_{\epsilon,k,k-1}(\mathbf{u})\Phi^{-1}_{\alpha,k}(\mathbf{u}). && (52)\end{aligned}$$

As shown in equation (52), the formula for $G_k(\mathbf{u})$ is not very useful because calculating the Kalman gain G_k requires to compute $\Phi_{\epsilon,k,k-1}(\mathbf{u})$ which is still unknown. In order to overcome this difficulty, we derive a recursive formula for $\Phi_{\epsilon,k,k-1}(\mathbf{u})$.

Consider the predicted state error at time frame $k+1$, $\epsilon_{k+1,k}(\mathbf{u})$,

$$\begin{aligned} \epsilon_{k+1,k}(\mathbf{u}) &= b_{k+1}(\mathbf{u}) - \hat{b}_{k+1|k}(\mathbf{u})\\ &= \gamma(\mathbf{u})b_k(\mathbf{u}) + \beta_k(\mathbf{u}) - \gamma(\mathbf{u})\hat{b}_{k|k-1}(\mathbf{u}) - G_k(\mathbf{u})\alpha_k(\mathbf{u})\\ &= \gamma(\mathbf{u})[b_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u})] - G_k(\mathbf{u})\alpha_k(\mathbf{u}) + \beta_k(\mathbf{u})\\ &= \gamma(\mathbf{u})[b_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u})] - G_k(\mathbf{u})[a_k(\mathbf{u}) - \hat{a}_{k|k-1}(\mathbf{u})] + \beta_k(\mathbf{u})\\ &= \gamma(\mathbf{u})[b_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u})] - G_k(\mathbf{u})[(b_k(\mathbf{u}) + n_k(\mathbf{u})) - \hat{b}_{k|k-1}(\mathbf{u})] + \beta_k(\mathbf{u})\\ &= \gamma(\mathbf{u})[\epsilon_{k,k-1}(\mathbf{u})] - G_k(\mathbf{u})\epsilon_{k,k-1}(\mathbf{u}) - G_k(\mathbf{u})n_k(\mathbf{u}) + \beta_k(\mathbf{u})\\ &= [\gamma(\mathbf{u}) - G_k(\mathbf{u})]\epsilon_{k,k-1}(\mathbf{u}) - G_k(\mathbf{u})n_k(\mathbf{u}) + \beta_k(\mathbf{u}).\end{aligned}$$

This implies that

$$\begin{aligned} \Phi_{\epsilon,k+1,k}(\mathbf{u}) &= E[\epsilon_{k+1,k}(\mathbf{u})\epsilon^H_{k+1,k}(\mathbf{u})]\\ &= [\gamma(\mathbf{u}) - G_k(\mathbf{u})]\Phi_{\epsilon,k,k-1}(\mathbf{u})[\gamma(\mathbf{u}) - G_k(\mathbf{u})]^H\\ &\quad + E[\beta_k(\mathbf{u})\beta^H_k(\mathbf{u})] + G_k(\mathbf{u})\Phi_n(\mathbf{u})G^H_k(\mathbf{u}).\end{aligned}$$

where $E[\beta_k(\mathbf{u})\beta^H_k(\mathbf{u})] = [1 - |\gamma(\mathbf{u})|^2]\Phi_b(\mathbf{u})$. If we expand the above equation, we obtain the called Riccati difference equation,

$$\begin{aligned} \Phi_{\epsilon,k+1,k}(\mathbf{u}) &= \gamma(\mathbf{u})\Phi_{\epsilon,k,k}(\mathbf{u})\gamma^H(\mathbf{u}) + E[\beta_k(\mathbf{u})\beta_k(\mathbf{u})^H]\\ &= \gamma(\mathbf{u})\Phi_{\epsilon,k,k}(\mathbf{u})\gamma^H(\mathbf{u}) + [1 - |\gamma(\mathbf{u})|^2]\Phi_b(\mathbf{u})\end{aligned}$$

where we denote $\Phi_{\epsilon,k,k}(\mathbf{u})$ by

$$\Phi_{\epsilon,k,k}(\mathbf{u}) = \Phi_{\epsilon,k,k-1}(\mathbf{u}) - [\gamma(\mathbf{u})]^{-1} G_k(\mathbf{u}) \Phi_{\epsilon,k,k-1}(\mathbf{u}) \tag{53}$$

In order to make the derived recursive formulas work, we impose the following initial conditions

$$\begin{aligned} \hat{b}_{1|0}(\mathbf{u}) &= 0 \\ \Phi_{\epsilon,1,0}(\mathbf{u}) &= E[b_1(\mathbf{u}) b_1^H(\mathbf{u})] = \Phi_b(\mathbf{u}) \end{aligned}$$

After deriving the desired recursive equations for our image model, we are ready to apply them to the problem addressed in Warren's work.

As mentioned previously, the importance of using innovation processes is the independency. In what follows we transform original processes considered in the image model to innovation processes and reformulate an equivalent hypothesis testing problem in terms of these innovation process $\alpha_k(\mathbf{u})$ and $\epsilon_{k,k-1}(\mathbf{u})$. As a result, the equivalent hypothesis testing problem can be described as follows.

$$\begin{aligned} H_0 : \quad & \alpha_k(\mathbf{u}) = a_k(\mathbf{u}) - \hat{a}_{k|k-1}(\mathbf{u}) \\ \textit{versus} \quad & \\ H_1 : \quad & \alpha_k(\mathbf{u}) = t_k(\mathbf{u}) + a_k(\mathbf{u}) - \hat{a}_{k|k-1}(\mathbf{u}) \end{aligned} \tag{54}$$

where $\alpha_k(\mathbf{u})$ is a white Gaussian process. The reason for $\alpha_k(\mathbf{u})$ being a white Gaussian process is that $\alpha_k(\mathbf{u})$ is independent and is a process resulting from a Gaussian process $a_k(\mathbf{u})$ from which the predicted estimate $\hat{a}_{k|k-1}(\mathbf{u})$ is subtracted which is also a linear combination of Gaussian processes.

The resulting log likelihood ratio test for this hypothesis testing problem is given by

$$LR_{\alpha(k)}(A, X_T) = \log \left[\frac{\mathbf{p}(\alpha(k)|H_1)}{\mathbf{p}(\alpha(k)|H_0)}\right]$$

where

$$\alpha(k) = [\alpha_1(\mathbf{u}), \cdots, \alpha_k(\mathbf{u})].$$

Since $\alpha_j(\mathbf{u})$s are independent, the log likelihood ratio test can be further expressed by

$$LR_{\alpha(k)}(A, X_T) = \log [\frac{\prod_{j=1}^{k} p(\alpha(k)|H_1)}{\prod_{j=1}^{k} p(\alpha(k)|H_0)}] = \sum_{j=1}^{k} \log [\frac{p(\alpha_j(\mathbf{u}|H_1)}{p(\alpha_j(\mathbf{u}|H_0)}] \quad (55)$$

$$= \frac{1}{2}\{\sum_{j=1}^{k} log [\frac{\Phi_{\alpha,j}(\mathbf{u}|H_1)}{\Phi_{\alpha,j}(\mathbf{u}|H_0)}] + \alpha_j(\mathbf{u}|H_0)\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_0)\alpha_j^H(\mathbf{u}|H_0)$$
$$-\alpha_j(\mathbf{u}|H_1)\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_1)\alpha_j^H(\mathbf{u}|H_1) \quad (56)$$

$$= \frac{1}{2}\{\sum_{j=1}^{k} log [\frac{\Phi_{\alpha,j}(\mathbf{u}|H_1)}{\Phi_{\alpha,j}(\mathbf{u}|H_0)}] + ||\alpha_j(\mathbf{u}|H_0)||^2_{\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_0)}$$
$$-||\alpha_j(\mathbf{u}|H_1)||^2_{\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_1)} \quad (57)$$

$$= LR_{\alpha(k-1)}(A, X_T) + \frac{1}{2}\{log [\frac{\Phi_{\alpha,k}(\mathbf{u}|H_1)}{\Phi_{\alpha,k}(\mathbf{u}|H_0)}]\}$$
$$+\frac{1}{2}\{||\alpha_k(\mathbf{u}|H_0)||^2_{\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_0)} - ||\alpha_k(\mathbf{u}|H_1)||^2_{\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_1)}\} \quad (58)$$

where

$$||\alpha_j(\mathbf{u}|H_i)||^2_{\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_i)} = \alpha_j(\mathbf{u}|H_i)\Phi_{\alpha,j}^{-1}(\mathbf{u}|H_i)\alpha_j^H(\mathbf{u}|H_i).$$

The last two equations (equations (57-58)) can be regarded as two parts. The first part is the likelihood ratio test obtained at time $k-1$ which is known at time k and the second part consisting of two terms which is made of new information obtained at time frame k under two hypotheses. The former represents previous and predicted information; the latter updates new information as time elapses. This fact is also demonstrated in equation (40) derived by Warren where $a_{k+1|1} - \hat{a}_{k+1|k,1}$ and $a_{k+1|0} - \hat{a}_{k+1|k,0}$ represent innovation processes generated by the observation processes $a_{k+1|0}$ and $a_{k+1|1}$ under hypothesis H_0 and H_1 respectively. The computation of the innovation process α_k and DFTs of its correlation matrix can be also recursively calculated by equations summaried as follows.

$$\alpha_k(\mathbf{u}) \equiv a_k(\mathbf{u}) - \hat{a}_{k|k-1}(\mathbf{u})$$

$$\epsilon_{k,k-1}(\mathbf{u}) \equiv b_k(\mathbf{u}) - \hat{b}_{k|k-1}(\mathbf{u})$$

$$\Phi_{\alpha,k}(\mathbf{u}) \equiv E[\alpha_k(\mathbf{u})\alpha_k^H(\mathbf{u})] \equiv \sum_{\mathbf{x}} \Lambda_{\alpha,k}(\mathbf{x})e^{-2\pi \mathbf{ux}}$$

$$\Phi_{\epsilon,k,k-1}(\mathbf{u}) \equiv E[\epsilon_{k,k-1}(\mathbf{u})\epsilon_{k,k-1}^H(\mathbf{u})] \equiv \sum_{\mathbf{x}} \Lambda_{\epsilon,k,k-1}(\mathbf{x})e^{-2\pi \mathbf{ux}}$$

$$
\begin{aligned}
\Phi_{\alpha,k}(\mathbf{u}) &= \Phi_{\epsilon,k,k-1}(\mathbf{u}) + \Phi_n(\mathbf{u}) \\
G_k(\mathbf{u}) &= \gamma(\mathbf{u})\Phi_{\epsilon,k,k-1}(\mathbf{u})\Phi^{-1}_{\alpha,k}(\mathbf{u}) \\
\hat{b}_{k+1|k}(\mathbf{u}) &= \gamma(\mathbf{u})\hat{b}_{k|k-1}(\mathbf{u}) + G_k(\mathbf{u})\alpha_k(\mathbf{u}) \\
\Phi_{\epsilon,k+1,k}(\mathbf{u}) &= \gamma(\mathbf{u})\Phi_{\epsilon,k,k}(\mathbf{u})\gamma^H(\mathbf{u}) + [1 - |\gamma(\mathbf{u})|^2]\Phi_b(\mathbf{u}) \\
\Phi_{\epsilon,k,k}(\mathbf{u}) &= \Phi_{\epsilon,k,k-1}(\mathbf{u}) - [\gamma(\mathbf{u})]^{-1}G_k(\mathbf{u})\Phi_{\epsilon,k,k-1}(\mathbf{u})
\end{aligned}
$$

with initial conditions given by

$$\hat{b}_{1|0}(\mathbf{u}) = 0; \quad LR_{\alpha(0)} = 0 \tag{59}$$

$$\Phi_{\epsilon,1,0}(\mathbf{u}) = E[b_1(\mathbf{u})b_1^H(\mathbf{u})] = \Phi_b(\mathbf{u}) \tag{60}$$

VI. Recursive Detection Algorithms for Multiple Spectral Bands

In order to extend Warren's algorithm to cover multiple spectral bands, we consider the two-spectral-band case. The extension to multiple bands can be done straightforwardly.

Two Spectral Bands

Let $\mathbf{x}$ and $\mathbf{y}$ be two sample random vectors producing spatial images and $a_k(\mathbf{u})$ and $a_k(\mathbf{v})$ are corresponding DFTs respectively where we use $\mathbf{u}, \mathbf{v}$ to distinguish two different spectral vectors. In general the pair of random vectors $(\mathbf{x}, \mathbf{y})$ is correlated, so is the pair $(\mathbf{u}, \mathbf{v})$. Let $a_k(\mathbf{u}, \mathbf{v})$ be the joint DFT of $I(\mathbf{x}, \mathbf{y})$ obtained based on the spatial image jointly generated by $(\mathbf{x}, \mathbf{y})$ in spatial domain. Then

$$a_k(\mathbf{u}, \mathbf{v}) = \sum_{\mathbf{x}} \sum_{\mathbf{y}} [I_k(\mathbf{x}, \mathbf{y})) - \overline{I}_0] e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})}.$$

where the notation dot "$\cdot$" denotes the inner product of two vectors.

Similarly, we can also define

$$\begin{aligned} b_k(\mathbf{u}, \mathbf{v}) &= \sum_{\mathbf{x}} \sum_{\mathbf{y}} [B_k(\mathbf{x}, \mathbf{y}) - \overline{B}] e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})} \\ n_k(\mathbf{u}, \mathbf{v}) &= \sum_{\mathbf{x}} \sum_{\mathbf{y}} [N_k(\mathbf{x}, \mathbf{y}) - \overline{N}] e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})} \\ t_k(\mathbf{u}, \mathbf{v}) &= \sum_{\mathbf{x}} \sum_{\mathbf{y}} [T_k(\mathbf{x}, \mathbf{y})] e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})} = A e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})} \mathbf{S}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

Extension of Warren's Recursive Algorithm

Following the same notations in Warren's work, we denote the covariance matrix of $[I_k(\mathbf{x}, \mathbf{y})]$ at time frame k by $\Phi_a(\mathbf{x}, \mathbf{y})_k$ and the DFT of $\Phi_a(\mathbf{x}, \mathbf{y})_k$ by $\Phi_a(\mathbf{u}, \mathbf{v})_k$ and related by

$$\begin{aligned} \Phi_a(\mathbf{u}, \mathbf{v})_k &= \sum_{\mathbf{x}} \sum_{\mathbf{y}} E\{[I_k(\mathbf{x}, \mathbf{y}) - \overline{I}_0][(I_k(\mathbf{x}', \mathbf{y}') - \overline{I}_0))^H]\} e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})} \qquad (61) \\ &= \sum_{\mathbf{x}} \sum_{\mathbf{y}} \Lambda_a(\mathbf{x}, \mathbf{y})_k e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})}. \qquad (62) \end{aligned}$$

Similarly we can define

$$\begin{aligned}\Phi_b(\mathbf{u},\mathbf{v})_k &= \sum_{\mathbf{x}}\sum_{\mathbf{y}} \Lambda_b(\mathbf{x},\mathbf{y}) e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})} \\ \Phi_n(\mathbf{u},\mathbf{v})_k &= \sum_{\mathbf{x}}\sum_{\mathbf{y}} C_n(\mathbf{x},\mathbf{y}) e^{-i2\pi(\mathbf{u},\mathbf{v})\cdot(\mathbf{x},\mathbf{y})},\end{aligned}$$

which yield

$$\Phi_a(\mathbf{u},\mathbf{v})_k = \Phi_b(\mathbf{u},\mathbf{v})_k + \Phi_n(\mathbf{u},\mathbf{v})_k \quad \text{under} \quad H_0; \tag{63}$$

$$\Phi_a(\mathbf{u},\mathbf{v})_k = t_k(\mathbf{u},\mathbf{v}) + \Phi_b(\mathbf{u},\mathbf{v})_k + \Phi_n(\mathbf{u},\mathbf{v})_k \quad \text{under} \quad H_1. \tag{64}$$

Moreover, the background and noise are assume to be statistically wide-sense homogeneous spatially and stationary temporally. So, the DFTs $\Phi_b(\mathbf{u},\mathbf{v})_k$ and $\Phi_n(\mathbf{u},\mathbf{v})_k$ at time k can be actually computed by $\Phi(\mathbf{u},\mathbf{v})_1$ and $\Gamma(\mathbf{u},\mathbf{v})$.

If we denote $\Phi(\mathbf{u},\mathbf{v})_1$ by $\Phi_b(\mathbf{u},\mathbf{v})$ and $[a_k(\mathbf{u},\mathbf{v}), a_{k-1}(\mathbf{u},\mathbf{v}), \cdots, a_1(\mathbf{u},\mathbf{v})]$ by $\mathbf{a}_k(\mathbf{u},\mathbf{v})$, Warren's recursive equations can be extended to cover two spectral bands as follows.

$$\begin{aligned}\Delta_0(\mathbf{u},\mathbf{v}) &\equiv \Phi_b(\mathbf{u},\mathbf{v}) + \Phi_n(\mathbf{u},\mathbf{v}) \\ \Gamma(\mathbf{u},\mathbf{v}) &\equiv \begin{bmatrix} \gamma_{11}(\mathbf{u},\mathbf{v}) & \gamma_{12}(\mathbf{u},\mathbf{v}) \\ \gamma_{21}(\mathbf{u},\mathbf{v}) & \gamma_{22}(\mathbf{u},\mathbf{v}) \end{bmatrix} \\ D_k(\mathbf{u},\mathbf{v}) &= \Phi_b(\mathbf{u},\mathbf{v})[\Gamma^*(\mathbf{u},\mathbf{v}), (\Gamma^2)^*(\mathbf{u},\mathbf{v}), \cdots, (\Gamma^k)^*(\mathbf{u},\mathbf{v})] \\ \mathbf{\Lambda_a}(k+1) &= \begin{bmatrix} \Sigma_b(\mathbf{u},\mathbf{v}) & D_k^*(\mathbf{u},\mathbf{v}) \\ D_k(\mathbf{u},\mathbf{v}) & \mathbf{\Lambda_a}(k) \end{bmatrix} \\ \Delta_k(\mathbf{u},\mathbf{v}) &= \Delta_0(\mathbf{u},\mathbf{v}) - \underline{D}_k^*(\mathbf{u},\mathbf{v})\mathbf{Y}_k(\mathbf{u},\mathbf{v}) \\ \mathbf{Y}_k(\mathbf{u},\mathbf{v}) &\equiv \mathbf{\Lambda_a}^{-1}(k)\underline{D}_k(\mathbf{u},\mathbf{v}) \\ \mathbf{\Lambda_a}^{-1}(k+1) &= \begin{bmatrix} \Delta_k^{-1}(\mathbf{u},\mathbf{v}) & -\Delta_k^{-1}(\mathbf{u},\mathbf{v})\mathbf{Y}_k^*(\mathbf{u},\mathbf{v}) \\ -\mathbf{Y}_k(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v}) & \mathbf{\Lambda_a}^{-1}(k)(\mathbf{u},\mathbf{v}) + \mathbf{Y}_k(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v})\mathbf{Y}_k^*(\mathbf{u},\mathbf{v}) \end{bmatrix} \\ \Delta_{k+1}(\mathbf{u},\mathbf{v}) &= \Phi_b(\mathbf{u},\mathbf{v})(1-|\gamma(\mathbf{u},\mathbf{v})|^2) + \Phi_n(\mathbf{u},\mathbf{v})(1+|\gamma(\mathbf{u},\mathbf{v})|^2) \\ &\quad -|\gamma(\mathbf{u},\mathbf{v})|^2\Phi_n(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v})\Phi_n(\mathbf{u},\mathbf{v}) \\ \mathbf{Y}_{k+1}^*(\mathbf{u},\mathbf{v}) &= \gamma^*(\mathbf{u},\mathbf{v})[1-\Phi_n(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v}), \Phi_n(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v})\mathbf{Y}_k^*(\mathbf{u},\mathbf{v})]\end{aligned}$$

It should be noted that the dimension of the matrix $\Phi_b(\mathbf{u},\mathbf{v})$ is $2N \times 2N$ whereas that of $\Phi_b(\mathbf{u})$ is $N \times N$.

Extension of Recursive Algorithm Using A Kalman Filter

One of nicest features of Kalman filtering approach is that the dimensionality of the state and observation models are arbitrary and can be any finite dimensions as long as they are appropriately aligned. This property makes an extension of the results for a single spectral band very easy and straightforward to multiple spectral bands. As assumed earlier, let $(\mathbf{x},\mathbf{y})$ be two joint random vectors which generates spatial images in spatial domain and let $(\mathbf{u},\mathbf{v})$ be the corresponding two spectral vectors used to describe their DFTs. The results for a single band can be easily extended as follows.

The desired two-spectral band log likelihood ratio test can be derived straightforwardly in terms of DFTs by using the same treatment done in Section V. Namely,

$$LR_{\alpha(k)}(A, X_T) = \log\,[\frac{\mathbf{p}(\alpha(k)|H_1)}{\mathbf{p}(\alpha(k)|H_0)}] \tag{65}$$

where

$$\alpha(k) = [\alpha_1(\mathbf{u},\mathbf{v}), \cdots, \alpha_k(\mathbf{u},\mathbf{v})],$$

and

$$\begin{aligned}
LR_{\alpha(k)}(A, X_T) &= \log\,[\frac{\prod_{j=1}^{k} p(\alpha(k)|H_1)}{\prod_{j=1}^{k} p(\alpha(k)|H_0)}] = \sum_{j=1}^{k} \log\,[\frac{p(\alpha_j(\mathbf{u},\mathbf{v}|H_1)}{p(\alpha_j(\mathbf{u},\mathbf{v}|H_0)}] && (66)\\
&= \frac{1}{2}\{\sum_{j=1}^{k} log\,[\frac{\Phi_{\alpha,j}(\mathbf{u},\mathbf{v}|H_1)}{\Phi_{\alpha,j}(\mathbf{u},\mathbf{v}|H_0)}]\alpha_j(\mathbf{u},\mathbf{v}|H_0)\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_0)\alpha_j^H(\mathbf{u},\mathbf{v}|H_0)\\
&\quad -\alpha_j(\mathbf{u},\mathbf{v}|H_1)\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_1)\alpha_j^H(\mathbf{u},\mathbf{v}|H_1) && (67)\\
&= \frac{1}{2}\{\sum_{j=1}^{k} log\,[\frac{\Phi_{\alpha,j}(\mathbf{u},\mathbf{v}|H_1)}{\Phi_{\alpha,j}(\mathbf{u},\mathbf{v}|H_0)}] + ||\alpha_j(\mathbf{u},\mathbf{v}|H_0)||^2_{\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_0)}\\
&\quad -||\alpha_j(\mathbf{u},\mathbf{v}|H_1)||^2_{\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_1)} && (68)\\
&= LR_{\alpha(k-1)}(A, X_T) + \frac{1}{2}\{log\,[\frac{\Phi_{\alpha,k}(\mathbf{u},\mathbf{v}|H_1)}{\Phi_{\alpha,k}(\mathbf{u},\mathbf{v}|H_0)}]\}\\
&\quad +\frac{1}{2}\{||\alpha_k(\mathbf{u},\mathbf{v}|H_0)||^2_{\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_0)} - ||\alpha_k(\mathbf{u},\mathbf{v}|H_1)||^2_{\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_1)}\} && (69)
\end{aligned}$$

where

$$||\alpha_j(\mathbf{u},\mathbf{v}|H_i)||^2_{\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_i)} = \alpha_j(\mathbf{u},\mathbf{v}|H_i)\Phi^{-1}_{\alpha,j}(\mathbf{u},\mathbf{v}|H_i)\alpha^H_j(\mathbf{u},\mathbf{v}|H_i).$$

and all the necessary recursive equations are summarized as follows.

$$\begin{aligned}
\alpha_k(\mathbf{u},\mathbf{v}) &\equiv a_k(\mathbf{u},\mathbf{v}) - \hat{a}_{k|k-1}(\mathbf{u},\mathbf{v}) \\
\epsilon_{k,k-1}(\mathbf{u},\mathbf{v}) &\equiv b_k(\mathbf{u},\mathbf{v}) - \hat{b}_{k|k-1}(\mathbf{u},\mathbf{v}) \\
\Phi_{\alpha,k}(\mathbf{u},\mathbf{v}) &\equiv E[\alpha_k(\mathbf{u},\mathbf{v})\alpha^H_k(\mathbf{u},\mathbf{v})] \equiv \sum_{\mathbf{x}}\sum_{\mathbf{y}} \Lambda_{\alpha,k}(\mathbf{x},\mathbf{y})e^{-2\pi \mathbf{u}\mathbf{x}}e^{-2\pi \mathbf{v}\mathbf{y}} \\
\Phi_{\epsilon,k,k-1}(\mathbf{u},\mathbf{v}) &\equiv E[\epsilon_{k,k-1}(\mathbf{u},\mathbf{v})\epsilon^H_{k,k-1}(\mathbf{u},\mathbf{v})] \equiv \sum_{\mathbf{x}}\sum_{\mathbf{y}} \Lambda_{\epsilon,k,k-1}(\mathbf{x},\mathbf{y})e^{-2\pi \mathbf{u}\mathbf{x}}e^{-2\pi \mathbf{v}\mathbf{y}} \\
\Phi_{\alpha,k}(\mathbf{u},\mathbf{v}) &= \Phi_{\epsilon,k,k-1}(\mathbf{u},\mathbf{v}) + \Phi_n(\mathbf{u},\mathbf{v}) \\
G_k(\mathbf{u},\mathbf{v}) &= \gamma(\mathbf{u},\mathbf{v})\Phi_{\epsilon,k,k-1}(\mathbf{u},\mathbf{v})\Phi^{-1}_{\alpha,k}(\mathbf{u},\mathbf{v}) \\
\hat{b}_{k+1|k}(\mathbf{u},\mathbf{v}) &= \gamma(\mathbf{u},\mathbf{v})\hat{b}_{k|k-1}(\mathbf{u},\mathbf{v}) + G_k(\mathbf{u},\mathbf{v})\alpha_k(\mathbf{u},\mathbf{v}) \\
\Phi_{\epsilon,k+1,k}(\mathbf{u},\mathbf{v}) &= \gamma(\mathbf{u},\mathbf{v})\Phi_{\epsilon,k,k}(\mathbf{u},\mathbf{v})\gamma^H(\mathbf{u},\mathbf{v}) + [1 - |\gamma(\mathbf{u},\mathbf{v})|^2]\Phi_b(\mathbf{u},\mathbf{v}) \\
\Phi_{\epsilon,k,k}(\mathbf{u},\mathbf{v}) &= \Phi_{\epsilon,k,k-1}(\mathbf{u},\mathbf{v}) - [\gamma(\mathbf{u},\mathbf{v})]^{-1}G_k(\mathbf{u},\mathbf{v})\Phi_{\epsilon,k,k-1}(\mathbf{u},\mathbf{v})
\end{aligned}$$

with initial conditions given by

$$\begin{aligned}
\hat{b}_{1|0}(\mathbf{u},\mathbf{v}) &= 0; \quad LR_{\alpha(0)}(A, X_T) = 0 \\
\Phi_{\epsilon,1,0}(\mathbf{u},\mathbf{v}) &= E[b_1(\mathbf{u},\mathbf{v})b^H_1(\mathbf{u},\mathbf{v})] = \Phi_b(\mathbf{u},\mathbf{v})
\end{aligned}$$

Separable Spectral Correlation

As we can see from the recursive equations derived for two-spectral band in both approaches, it is very difficult to calculate correlation matrices for the background and noise, particularly, their inverses because the dimensionality increases by 4 and correlations between spatial images in two spectral bands must be taken into account. One way to reduce computational complexity is to assume that all the power spectra obtained from different spectral bands are separable. In other words, under this circumstance, the observation process is

spectral-wise (not sample-wise) correlated. Namely, if we let $[a_1(\mathbf{u},\mathbf{v})] = [a_1(\mathbf{u}), a_1(\mathbf{v})]$ and only correlations between sample images $a_1(\mathbf{u})$ from one spectral band and sample images $a_1(\mathbf{v})$ from the other will be considered. By doing so the power spectrum of the observation process $[a_1(\mathbf{u}), a_1(\mathbf{v})]$, $\Phi_a(\mathbf{u},\mathbf{v})_k$ can be much simplified and given by

$$\Phi_a(\mathbf{u},\mathbf{v})_k = \begin{bmatrix} E[a_k(\mathbf{u})(a_k(\mathbf{u}))^H] & E[a_k(\mathbf{u})(a_k(\mathbf{v}))^H] \\ E[a_k(\mathbf{v})(a_k(\mathbf{u}))^H] & E[a_k(\mathbf{v})(a_k(\mathbf{v}))^H] \end{bmatrix}, \tag{70}$$

where

$$\begin{aligned} E[a_k(\mathbf{u})(a_k(\mathbf{u}))^H] &= \sum_{\mathbf{x}}\sum_{\mathbf{y}} \Lambda_a(\mathbf{x})\Lambda_b(\mathbf{x}')e^{-i2\pi\mathbf{u}\cdot\mathbf{x}}e^{i2\pi\mathbf{v}\cdot\mathbf{x}'} \\ E[a_k(\mathbf{u})(a_k(\mathbf{v}))^H] &= \sum_{\mathbf{x}}\sum_{\mathbf{y}} \Lambda_a(\mathbf{x})\Lambda_b(\mathbf{y})e^{-i2\pi\mathbf{u}\cdot\mathbf{x}}e^{i2\pi\mathbf{v}\cdot\mathbf{y}} \\ E[a_k(\mathbf{v})(a_k(\mathbf{u}))^H] &= \sum_{\mathbf{y}}\sum_{\mathbf{x}} \Lambda_a(\mathbf{y})\Lambda_b(\mathbf{x})e^{-i2\pi\mathbf{u}\cdot\mathbf{y}}e^{i2\pi\mathbf{v}\cdot\mathbf{x}} \\ E[a_k(\mathbf{v})(a_k(\mathbf{v}))^H] &= \sum_{\mathbf{y}}\sum_{\mathbf{y}'} \Lambda_a(\mathbf{y})\Lambda_b(\mathbf{y}')e^{-i2\pi\mathbf{u}\cdot\mathbf{y}}e^{i2\pi\mathbf{v}\cdot\mathbf{y}'} \end{aligned}$$

Therefore,

$$\begin{aligned} \Phi_b(\mathbf{u},\mathbf{v}) &= \sum_{\mathbf{x}}\sum_{\mathbf{y}} \Lambda_b(\mathbf{x})\Lambda_b(\mathbf{y})e^{-i2\pi\mathbf{u}\cdot\mathbf{x}}e^{-i2\pi\mathbf{v}\cdot\mathbf{y}} \\ \Phi_n(\mathbf{u},\mathbf{v}) &= \sum_{\mathbf{x}}\sum_{\mathbf{y}} C_n(\mathbf{x})C_n(\mathbf{y})e^{-i2\pi\mathbf{u}\cdot\mathbf{x}}e^{-i2\pi\mathbf{v}\cdot\mathbf{y}} \end{aligned}$$

If we denote $\Phi_1(\mathbf{u},\mathbf{v})$ by $\Phi_b(\mathbf{u},\mathbf{v})$ and $[a_k(\mathbf{u}), a_k(\mathbf{v}), a_{k-1}(\mathbf{u}), a_{k-1}(\mathbf{v}), \cdots, a_1(\mathbf{u}), a_1(\mathbf{v})]$ by $\mathbf{a}_k(\mathbf{u},\mathbf{v})$, then

$$\begin{aligned} \Delta_0(\mathbf{u},\mathbf{v}) &\equiv \Phi_b(\mathbf{u},\mathbf{v}) + \Phi_n(\mathbf{u},\mathbf{v}) \\ \Gamma(\mathbf{u},\mathbf{v}) &\equiv \begin{bmatrix} \gamma(\mathbf{u}) & 0 \\ 0 & \gamma(\mathbf{v}) \end{bmatrix} \\ D_k(\mathbf{u},\mathbf{v}) &= \Phi_b(\mathbf{u},\mathbf{v})[\Gamma^*(\mathbf{u},\mathbf{v}), (\Gamma^2)^*(\mathbf{u},\mathbf{v}), \cdots, (\Gamma^k)^*(\mathbf{u},\mathbf{v})] \\ \mathbf{\Lambda_a}(k+1) &= \begin{bmatrix} \Sigma_b(\mathbf{u},\mathbf{v}) & D_k^*(\mathbf{u},\mathbf{v}) \\ D_k(\mathbf{u},\mathbf{v}) & \mathbf{\Lambda_a}(k) \end{bmatrix} \\ \Delta_k(\mathbf{u},\mathbf{v}) &= \Delta_0(\mathbf{u},\mathbf{v}) - \underline{D}_k^*(\mathbf{u},\mathbf{v})\mathbf{Y}_k(\mathbf{u},\mathbf{v}) \\ \mathbf{Y}_k(\mathbf{u},\mathbf{v}) &\equiv \mathbf{\Lambda_a}^{-1}(k)\underline{D}_k(\mathbf{u},\mathbf{v}) \end{aligned}$$

$$\begin{aligned}\Delta_{k+1}(\mathbf{u},\mathbf{v}) &= \Phi_b(\mathbf{u},\mathbf{v})(1-|\gamma(\mathbf{u},\mathbf{v})|^2)+\Phi_n(\mathbf{u},\mathbf{v})(1+|\gamma(\mathbf{u},\mathbf{v})|^2)\\ &\quad -|\gamma(\mathbf{u},\mathbf{v})|^2\Phi_n(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v})\Phi_n(\mathbf{u},\mathbf{v})\\ \mathbf{Y}_{k+1}^*(\mathbf{u},\mathbf{v}) &= \gamma^*(\mathbf{u},\mathbf{v})[1-\Phi_n(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v}),\Phi_n(\mathbf{u},\mathbf{v})\Delta_k^{-1}(\mathbf{u},\mathbf{v})\mathbf{Y}_k^*(\mathbf{u},\mathbf{v})]\end{aligned}$$

Separable Spectral and Markov Spatial Correlation

Thus far, all derivations given above assume that DFTs of spatial images obtained from each band are spatially correlated. In multiple-band case correlations exist not only between spatial sample images but also between time frames. In the section of separable spectral correlation we considered the case that sample images are spatially correlated in each band and temporally blockwise correlated. In this section we can further reduce complexity by assuming that the spatial correlation has Markov property. Since the processes $\mathbf{x}$ and $\mathbf{y}$ considered in the image model are generally Gaussian, these observation processes can be further described by Gaussian-Markov processes. One of most important properties of a Gaussian-Markov process is that its covariance matrix can be characterized by powers of a constant with abosulte value less than 1.

More precisely, let the covariance matrix of the random vector $\mathbf{x}$ be $\Lambda(\mathbf{x})$, then the entry in the position (j,l) of $\Lambda(\mathbf{x})$ can be expressed by $\rho^{|j-l|}$ for some constant $|\rho|<1$. Applying this property to the covariance matrix of the background $B_k(\mathbf{x})$, $\Lambda_B(\mathbf{x})_k$ yields

$$\Lambda_B(\mathbf{x})_k = [\rho_1^{|j-l|}\rho_2^{|m-n|}],$$

where $\rho_1^{|j-l|}$ describes the horizontal spatial correlation between $x_{j,m}$ and $x_{l,m}$ and $\rho_2^{|m-n|}$ indicates the vertical spatial correlation between $x_{j,m}$ and $x_{j,n}$.

Furthermore, using the property of spectral separability we can derive the correlation matrix of two sample images which generate power spectra in two different bands as follows.

$$\Lambda_B(\mathbf{x},\mathbf{y})_k = [\rho_1^{|j-l|}\rho_2^{|m-n|}\sigma_1^{|r-s|}\sigma_2^{|t-w|}]. \tag{71}$$

where ρ_1, σ_1 and ρ_2, σ_2 are horizontal and vertical spatial correlations of images $\mathbf{x}$ and $\mathbf{y}$ respectively. Substituting equation (71) for $\Lambda_B(\mathbf{x}, \mathbf{y})_k$ in the recursive equations derived for separable spectral correlation will yield a set of similar necessary recursive formulas.

VII. Conclusion

The main theme in this report is to present an alternative to develop a new recursive algorithm for thermal image detection and extend Warren's and new algorithms from a single spectral band to multiple spectral bands. The idea to derive the new algorithm is Kalman filter theory. This approach is more intuitive than Warren's formulation. A complete mathematical derivation is given. There are some future directions needed to be investigated. (1) Adaptive techniques should be considered to improve effectiveness of suppressing background clutter. (2) In practical situations the probability laws governing the background clutter are generally not informed. We must estimate them by virtue of sample images. For instance, for Gaussian processes the mean and variance matrices must be estimated before applying recursive algorithms. (3) For real time implementation, computer simulation for results generated by this report needs to be done for synthetic images. (4) Development of systolic architectures for implementing thermal image detectors for parallel processing is needed to speed up computational rates. Finally, (5) An unconventional approach, neural network should be investigated to increase the capability of thermal image detectors to adapt the various environments.

References

1. R.E. Warren, "Algorithm for image enhancement and detection of chemical vapor using a thermal imager," Final Report, US Army Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, MD, 1987.

2. L. Carr, D. McPherrin and R.E. Warren, "Low contrast detection algorithm," Technical Report, US Army Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, MD, 1990.

3. H.E. Rauch, W.I. Futterman and D.B. Kemmer, "Background suppression and tracking with a staring," *Optical Engineering*, Jan/Feb 1981, Vol. 20, No. 1, pp. 103-110.

4. A. Aridgides, G. Cook and S.M..K. Zonca, "Correlated background adaptive clutter suppression and normalization techniques," *SPIE, Multispectral Image Processing and Enhancement*, Vol. 933, 1988, pp. 32-44